

Perceptual Quality-Regulable Video Coding System With Region-Based Rate Control Scheme

Guan-Lin Wu, Yu-Jie Fu, Sheng-Chieh Huang, and Shao-Yi Chien, *Member, IEEE*

Abstract—In this paper, we discuss a region-based perceptual quality-regulable H.264 video encoder system that we developed. The ability to adjust the quality of specific regions of a source video to a predefined level of quality is an essential technique for region-based video applications. We use the structural similarity index as the quality metric for distortion-quantization modeling and develop a bit allocation and rate control scheme for enhancing regional perceptual quality. Exploiting the relationship between the reconstructed macroblock and the best predicted macroblock from mode decision, a novel quantization parameter prediction method is built and used to achieve the target video quality of the processed macroblock. Experimental results show that the system model has only 0.013 quality error in average. Moreover, the proposed region-based rate control system can encode video well under a bitrate constraint with a 0.1% bitrate error in average. For the situation of the low bitrate constraint, the proposed system can encode video with a 0.5% bit error rate in average and enhance the quality of the target regions.

Index Terms—H.264, perceptual video coding, structural similarity (SSIM).

I. INTRODUCTION

REGION-based video coding is used to create a better visual experience for users by enhancing the quality of a specific area of the source video through resource management, where the resource can be channel bandwidth or computational power. Most of video coding applications give equal importance to every macroblock (MB) regardless of its relative importance to users. However, for many video applications, users pay more attention to the regions of interest (ROI) in a video source. A region-based video coding scheme is especially useful for personal devices, such as cell phones and PDAs, particularly for real-time video communication. Therefore, it is desired to develop a region-based video coding system that considers the tradeoff between the available resources and the video quality.

Rate control is a critical technique in video coding system to allocate bitrate resources [1]. It tries to use the available

channel bandwidth and coding buffer size to achieve the best video quality. The main process in rate control with a rate constraint is to decide the idea encoding rate for a source video by properly choosing a sequence of quantization parameters. The relationship between the coding rate and the quantization parameter (QP) can be characterized as a rate-quantization (R-Q) model which decides the QP value for each source video given a target bitrate [2]–[9]. To achieve the better rate-distortion optimization, a distortion-quantization (D-Q) model, which models the relationship between distortion and QP, is adopted to allocate the bit budget for different video sources [7], [8]. These schemes use conventional objective distortion metrics, such as the mean squared error (MSE) or the mean absolute difference (MAD). However, these objective distortion metrics have been shown to have low correlation with the human visual system (HVS) [10]–[13].

Many schemes have been proposed to incorporate the characteristics of the human visual system into rate control [14]–[24]. In [14], the quantization step of each MB is determined with the rate-distortion model considering the local perceptual cues in the video signal, such as luminance adaptation, texture masking, and skin color. In [17], they use the direct frame difference and skin-tone information as the weighting factor for determining the QP of each MB. The authors in [16] have proposed a visual distortion sensitivity model to indicate the regions where the distortion could be tolerated and therefore fewer bits could be allocated. The model exploits the non-uniform spatiotemporal sensitivity characteristics of the HVS based on motion, textural structures, motion attention, spatial velocity visual sensitivity and visual masking. In [18], the foveation model and the just-noticeable-distortion (JND) model have been combined into a unified foveated JND model for the QP adjustment and the Lagrange multiplier in the rate-distortion optimization scheme. In [19] and [20], attention model, JND, contrast sensitivity function, and the structural similarity (SSIM) index have been involved in the perceptual model that is used to adjust the QP of each MB in video. However, these methods use some complicated human perceptual factors as the weight of each region to enhance the perceptual quality. They also use heuristic weighting methods to adjust the QP for bitrate allocation.

Several video or image quality metrics that consider the characteristics of HVS have been developed recently [10]–[13] to better quality evaluate processed signals. These metrics correlate with human perception higher than MSE or MAD [25]. Among the perceptual metrics proposed in the literature, SSIM [10] has been shown to be effective and computationally efficient. A number of image/video quality

Manuscript received August 20, 2012; revised January 21, 2013; accepted February 3, 2013. Date of publication February 11, 2013; date of current version April 12, 2013. This work was supported by the National Science Council, under Grant NSC 100-2221-E-002-090-MY3 and NSC 101-2220-E-002-010-MY3. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Joan Serra-Sagrista.

G.-L. Wu, Y.-J. Fu, and S.-Y. Chien are with Media IC and System Laboratory, Graduate Institute of Electronics Engineering and Department of Electrical Engineering, National Taiwan University, Taipei 10617, Taiwan (e-mail: vince.wuguanlin@gmail.com; fmichael8000@hotmail.com; sychien@cc.ee.ntu.edu.tw).

S.-C. Huang is with the Department of Electrical Engineering, National Chiao-Tung University, Hsinchu 30010, Taiwan (e-mail: schuang@cn.nctu.edu.tw).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2013.2247409



Fig. 1. Basic concept of the quality-regulable coding system. The quality of ROI regions can be adjusted based on user's target quality under the constraint of the same bitrate. (a) ROI (white region) and non-ROI (black region). (b) ROI quality:50 and non-ROI quality:70 Bitrate:100. (c) ROI quality:60 and non-ROI quality:60 Bitrate:100. (d) ROI quality:70 and non-ROI quality:50 Bitrate:100.

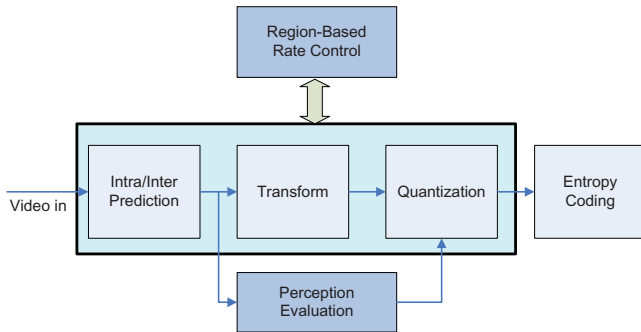


Fig. 2. Block diagram of the proposed region-based perceptual quality-regulable video coding scheme.

assessment methods using the SSIM index have been developed and applied to image/video processing and compression [20], [26]–[29]. Optimum bit allocation for image compression based on the SSIM index has been proposed in [26] and [27]. In [21], the authors used the SSIM index as the quality metric for rate-distortion modeling, and developed an optimum rate control algorithm for video coding.

We address region-based video coding schemes, which were overlooked by the papers mentioned above. Although previous works use a perceptual model to determine where to enhance the quality of the video, they do not provide the quality-regulable capability, which is the ability to adjust the video quality of target regions based on the user's preference. In some video communication applications, such as video conferencing, and in sports videos, users might want to focus on a specific region in a video sequence. Moreover, users might want the target region to be displayed at a specific level of quality. At the same time, the user's preferred quality for the target region might extend beyond the bitrate constraint, which depends on the channel bandwidth. In this situation, a well-designed rate control scheme has to allocate the bitrate and adjust the QP to meet the bitrate constraint. In this paper, we built a new predictive QP estimation model based on perceptual quality metrics and used it to regulate the video quality according to a target perceptual quality. Given a target bit-rate, the proposed algorithm tries to achieve the target visual quality on areas where the human viewer might focus. Moreover, in instances where the bit budget is insufficient to achieve the target quality, our scheme can also perform smarter bit rate allocations between target and non-target regions to achieve better perceptual quality for target regions.

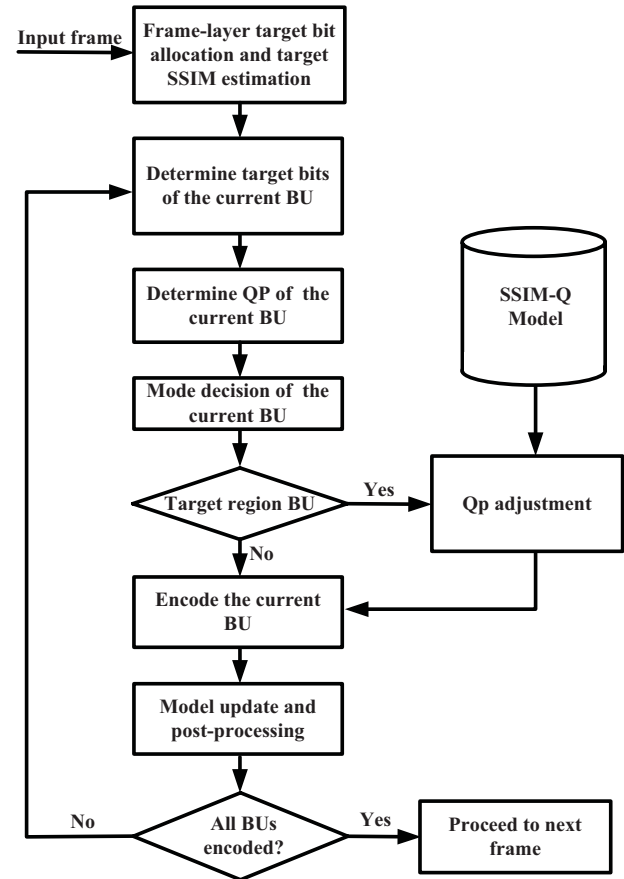


Fig. 3. System flow of the proposed region-based quality-regulable video coding scheme.

The organization of this paper is as follows: Section II describes the proposed system with H.264 encoder. Section III shows the evaluation system with subjective experiments for the proposed system. The experimental results and data are also analyzed. Finally, a brief conclusion is given in Section IV.

II. PROPOSED SYSTEM

Fig. 1 shows the goal of the proposed region-based perceptual quality-regulable video coding scheme. Under the constraint of using the same bitrate for encoding the same source video, the quality of target regions can be adjusted by users. Fig. 2 shows the basic system of our proposed video

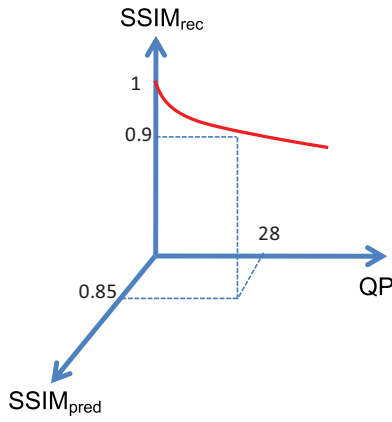


Fig. 4. SSIM-Q modeling concept.

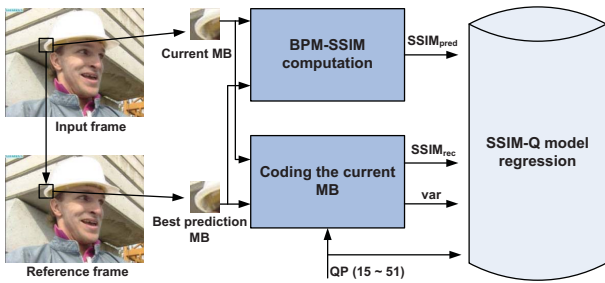


Fig. 5. SSIM-Q modeling procedures.

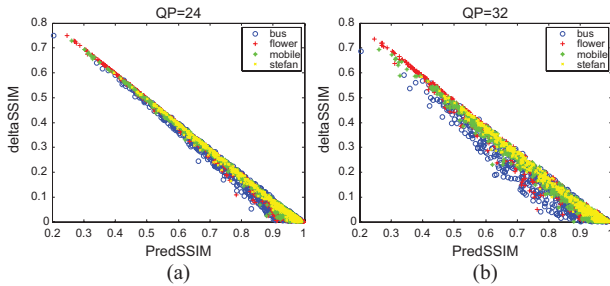


Fig. 6. $SSIM_{pred}$ (x-coordinate) versus $\Delta SSIM_{rec}$ (y-coordinate) with various QPs. (a) QP = 24. (b) QP = 32.

coding scheme within the video encoding flow, as well as the block diagrams of Intra/Inter Prediction, Transform, Quantization, and Entropy Coding, which are the required encoding modules. To achieve perceptual quality-regulability, we developed the Perception Evaluation module and the Region-Based Rate Control module in our system. The Perception Evaluation module calculates the coding and perceptual information of the current MB for the Quantization stage to adjust the QP of the current MB. The Region-Based Rate Control module performs the overall system control for bit allocation and QP adjustment to achieve the target perceptual quality based on coding and perceptual information. Note that when the channel bitrate is not sufficient to achieve the target quality for the target regions, the Region-Based Rate Control scheme also needs to control the bit allocation in order to avoid exceeding the bit budget and thereby, resulting in unstable quality as shown in the following source video.

Fig. 3 shows the detailed system flow of the proposed video coding system. In order to achieve the target perceptual quality

TABLE I
 R^2 STATISTICS OF REGRESSION RESULTS FOR CIF FORMAT SEQUENCES

Sequence	Bus	Flower	Football	Mobile	Stefan
R^2	0.9607	0.9837	0.9254	0.9808	0.9876

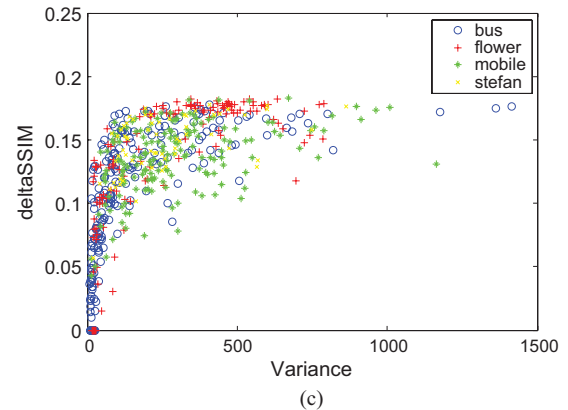
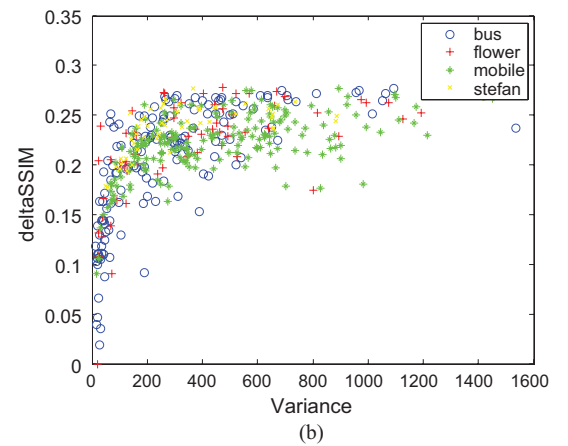
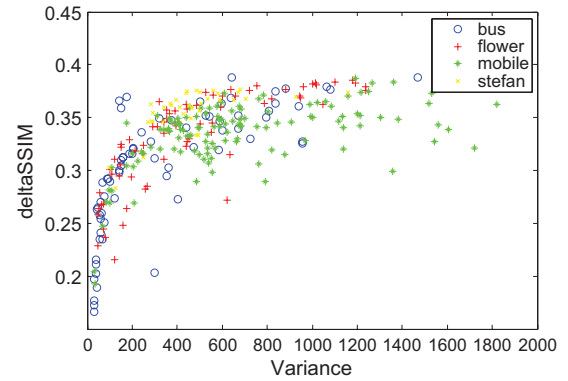


Fig. 7. Residual variance (x-coordinate) versus $\Delta SSIM_{rec}$ (y-coordinate) with fixed QP and $SSIM_{pred}$. (a) QP = 35, $SSIM_{pred} = 0.60$. (b) QP = 35, $SSIM_{pred} = 0.71$. (c) QP = 35, $SSIM_{pred} = 0.81$.

in the target region, we had to define the quality metrics for the system. Due to the required level of effectiveness and computational efficiency of the quality evaluation, the SSIM index was adopted in this paper. To determine the relationship between perceptual quality and QP, a perceptual metric and quantization parameter model (SSIM-Q) using a SSIM index had to be developed. Based on the SSIM-Q model, we can adjust the perceptual quality of the target region to a target SSIM value by adjusting the QP of the processed basic

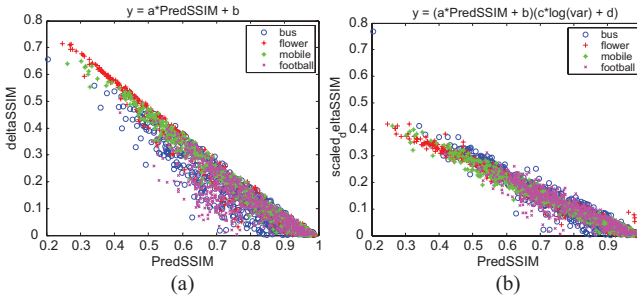


Fig. 8. Comparison of regression results (a) without residual variance consideration and (b) with residual variance consideration.

unit (BU). First, the frame layer bit allocation scheme tries to allocate the appropriate bit budget to the current frame. Thereafter, the appropriate bit budgets for non-target and target regions are estimated and allocated. To avoid exceeding the bit budget, we estimate the SSIM value that can be achieved using the target bitrate for the target region. If the target SSIM defined by users for the target region is smaller than the estimated SSIM, it means the target SSIM can be achieved within the bit budget. Otherwise, the estimated SSIM becomes the new target SSIM in the following coding flow. Next, the target number of bits is then estimated and allocated for each BU, depending on the bit budget for the region type (target or non-target). After that, the QP for each BU is computed based on the R-Q model. Then, the rate-distortion optimization mode decision process is performed to obtain the best prediction MB. The QP for the MBs in the target region is adjusted to achieve the target SSIM quality based on the SSIM-Q model. Finally, transform, quantization and entropy coding are performed to encode the current BU. The details of each main functional block are described in the following sections.

A. Structural Similarity Index as Quality Metric

SSIM index was adopted to help the coding system measure the video quality, because the sum of absolute difference (SAD) and the sum of squared difference (SSD) are poorly correlated with human perception [10]. The idea of SSIM is that the human visual system (HVS) is highly adapted to extract structural information from the image. The SSIM index measures the luminance similarity, contrast similarity, and structural similarity between two images block by block. The SSIM index is defined as follows:

$$SSIM(\mathbf{x}, \mathbf{y}) = l(\mathbf{x}, \mathbf{y}) \cdot c(\mathbf{x}, \mathbf{y}) \cdot s(\mathbf{x}, \mathbf{y}) \quad (1)$$

$$l(\mathbf{x}, \mathbf{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (2)$$

$$c(\mathbf{x}, \mathbf{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (3)$$

$$s(\mathbf{x}, \mathbf{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (4)$$

where \mathbf{x} and \mathbf{y} are two image blocks, μ_x and μ_y are the means of \mathbf{x} and \mathbf{y} , σ_x^2 and σ_y^2 are the variance of \mathbf{x} and \mathbf{y} , σ_{xy} is the cross variance between \mathbf{x} and \mathbf{y} , and C_1 , C_2 , and C_3 are three constants introduced in [10].

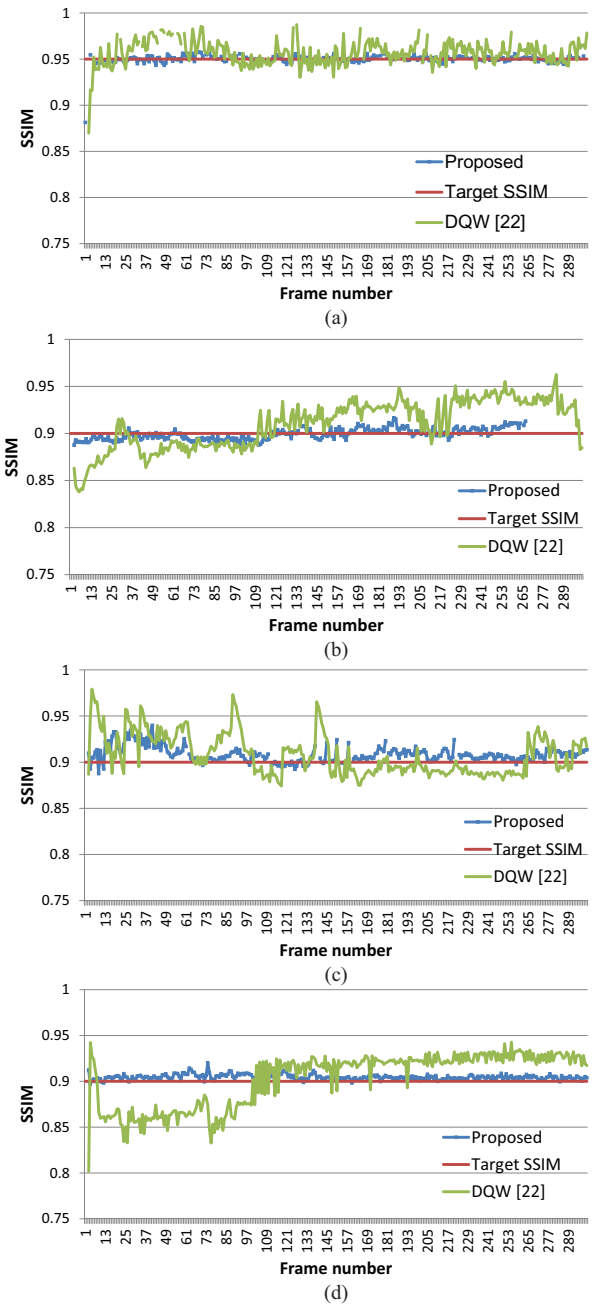


Fig. 9. Experimental results of the proposed system with the difference between the target SSIMs and the actual SSIM values for the target regions frame-by-frame. (a) *News*. (b) *Paris*. (c) *Crew*. (d) *Coastguard*.

B. SSIM-Q Model

Fig. 4 shows the basic concept of the proposed SSIM-Q model. $SSIM_{rec}$ is the SSIM value of the final reconstructed MB. It represents the quality of the final encoded MB. $SSIM_{pred}$ is the SSIM value between the current MB and the best predicted MB, which is determined after the inter/intra mode decision, and which is the lower bound of the SSIM value of the final reconstructed MB. As more residual data is preserved by decreasing the QP for the current MB, $SSIM_{rec}$ becomes larger and the perceptual quality improves.

Statistical analysis is adopted to obtain the relationship between $SSIM_{rec}$, $SSIM_{pred}$, and the corresponding QP.

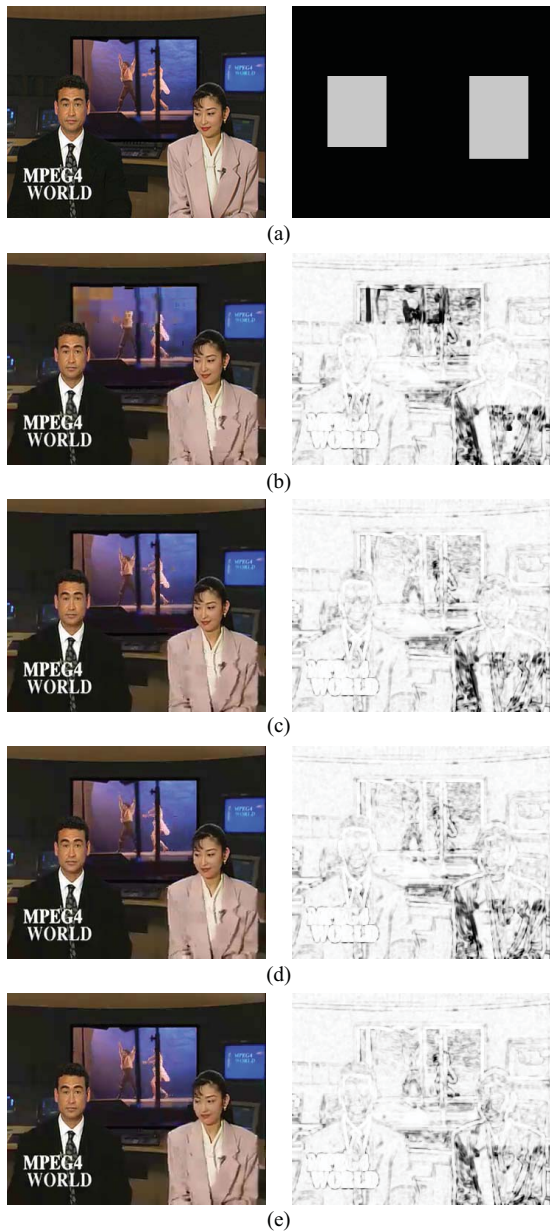


Fig. 10. Subjective demonstration of performance of the proposed system. The left side is the 200th frame of the reconstructed sequence “News.” The right side is the corresponding SSIM map. (a) Original frame and the target region (white region). (b) Target SSIM:0.95 ROI SSIM:0.9487. (c) Target SSIM:0.90 ROI SSIM:0.8932. (d) Target SSIM:0.85 ROI SSIM:0.8366. (e) Target SSIM:0.80 ROI SSIM:0.7844.

Fig. 5 illustrates the SSIM-Q modeling procedures. The “BPM-SSIM Computation” stage computes the SSIM index between the current MB and the best prediction MB (denoted as $SSIM_{pred}$) according to (1). The information about $SSIM_{rec}$, $SSIM_{pred}$, and QP (from 15 to 51 in this paper) are gathered after standard video encoding. By statistical analysis, we want to obtain the relationship in the form of

$$SSIM_{rec} = f(QP, SSIM_{pred}). \quad (5)$$

Next, (5) is transferred into the following form:

$$\Delta SSIM_{rec} = SSIM_{rec} - SSIM_{pred} = g(QP, SSIM_{pred}). \quad (6)$$

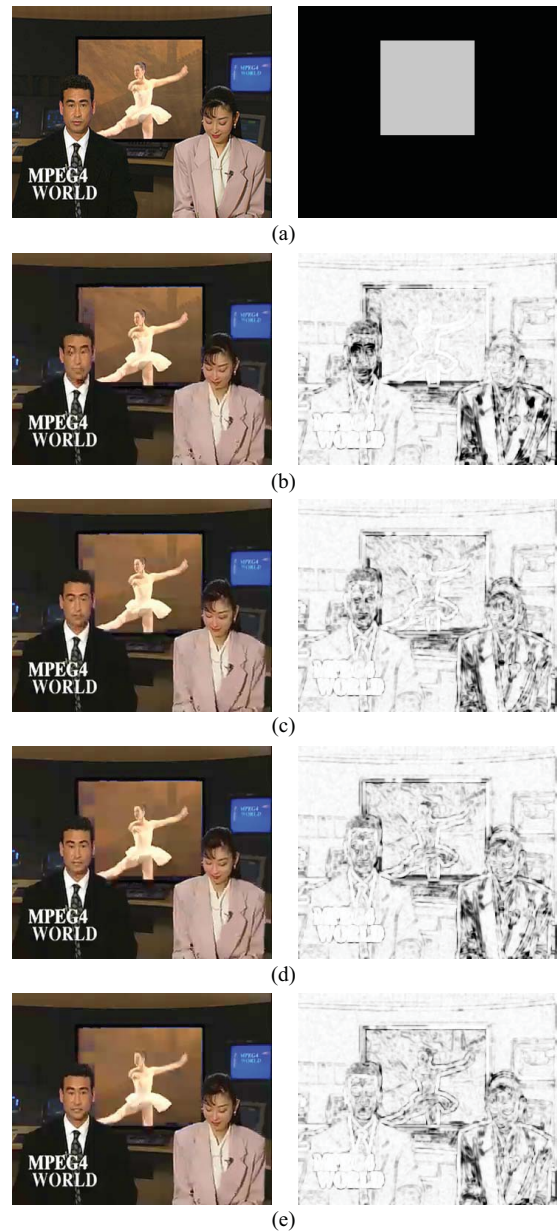


Fig. 11. Subjective demonstration of performance of the proposed system. The left side is the 91th frame of the reconstructed sequence “News.” The right side is the corresponding SSIM map. (a) Original frame and the target region (white region). (b) Target SSIM:0.95 ROI SSIM:0.9596. (c) Target SSIM:0.90 ROI SSIM:0.9023. (d) Target SSIM:0.85 ROI SSIM:0.8531. (e) Target SSIM:0.80 ROI SSIM:0.7945.

The variable $SSIM_{rec}$ is changed to $\Delta SSIM_{rec}$ because $SSIM_{rec}$ is larger or equal to $SSIM_{pred}$. Thus we only need to know the increment from $SSIM_{pred}$ to $SSIM_{rec}$.

By observing the $SSIM_{pred}$ and $\Delta SSIM_{rec}$ statistics with various QPs shown in Fig. 8 (a), we find a fitting relationship between $SSIM_{pred}$ and $\Delta SSIM_{rec}$ for a given QP. Thus we fit the data points by the following:

$$\Delta SSIM_{rec} = a \cdot SSIM_{pred} + b \cdot QP + c \cdot SSIM_{pred} \cdot QP + d \quad (7)$$

where a , b , c , and d are model parameters.

We further observe the regression results shown in Fig. 6. As the QP gets larger, the data points become more scattered

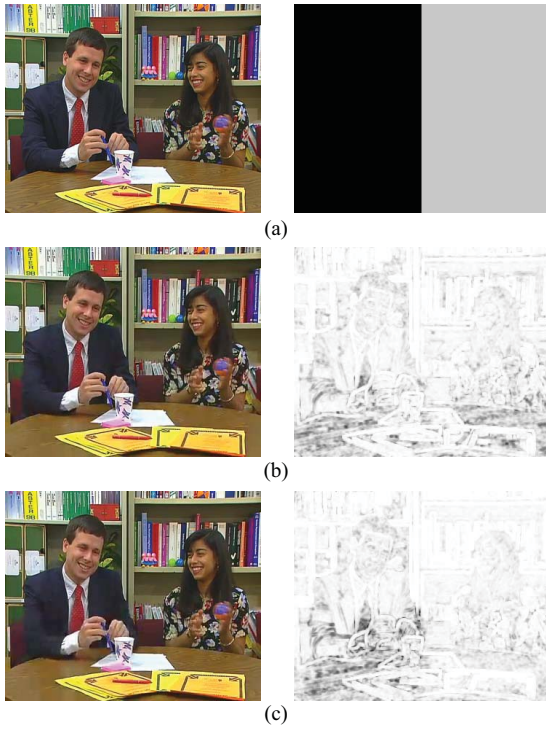


Fig. 12. Subjective demonstration of performance of the proposed system. The left side is the 150th frame of the reconstructed sequence “paris.” The right side is the corresponding SSIM map. (a) Original frame and the target region map (white region). (b) JM RC Bitrate:12144 ROI SSIM:0.9351. (c) Proposed Bitrate:12704 ROI SSIM:0.9601.

and more unlike a straight-line distribution. That is, when the QP gets larger, the idea $\Delta SSIM_{rec}$ can not be well fitted by only using QP and $SSIM_{pred}$. Based on our simulation and exploration, we found that given QP and $SSIM_{pred}$, there is a logarithmic relationship between residual variance and $\Delta SSIM_{rec}$. Fig. 7 shows the experimental results of this relationship using fixed QP. Fig. 8(b) shows the data points becoming more centralized when considering the residual variance. Finally, the modified regression mode from the (7) model is shown as follows:

$$\Delta SSIM_{rec} = (a \cdot SSIM_{pred} + b \cdot QP + c \cdot SSIM_{pred} \cdot QP + d) \cdot (e \cdot \log(var) + f) \quad (8)$$

where “var” denotes residual variance which is obtained from the difference between the best predicted MB and the current MB. a , b , c , d , e , and f are model parameters. R^2 [30] is adopted as a metric to evaluate the goodness of regression fitting. A higher R^2 value implies a better fitting, and the maximum R^2 value is 1, which occurs when all the data points are fitted perfectly. The R^2 statistics of the regression results with six test sequences are shown in Table I. By evaluating the same set of fitting coefficients for different sequences, we see that all the R^2 values of the fitting for different sequences are close to 1.

C. Region-Based Frame-Layer Target Bit Allocation and Target SSIM Estimation

Our focus is on the bit allocation scheme at the frame-layer and BU layer. Therefore, the schemes described in

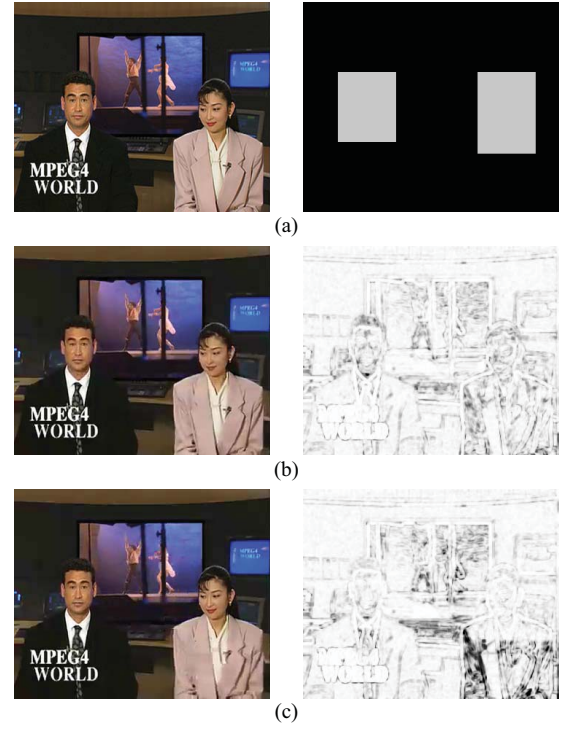


Fig. 13. Subjective demonstration of performance of the proposed system. The left side is the 200th frame of the reconstructed sequence “news.” The right side is the corresponding SSIM map. (a) Original frame and the target region map (white region). (b) JM RC Bitrate:2792 ROI SSIM:0.8946. (c) Proposed Bitrate:2064 ROI SSIM:0.9410.

JVT-G012 [6] were adopted for GOP-layer bit allocation. Assume Ψ and Ω are the target region and non-target region in one frame respectively. T_i is the target bit budget for the i th frame-by-frame layer rate control, and T_i can be obtained based on the method described in JVT-G012. Then, we assign the bit budget T_i^Ω for the non-target region in the i th frame as follows

$$T_i^\Omega = \frac{MAD_i^\Omega}{MAD_i^\Psi + MAD_i^\Omega} \cdot \frac{R}{f} \cdot \frac{1}{M} \quad (9)$$

where R is the available channel bandwidth; f denotes the frame rate; MAD_i^Ψ and MAD_i^Ω are the total MAD values of all the BUs of the target region and the non-target region in the i th frame, respectively. Similar to the method in JVT-G012, we predict the MAD of the current BU in our scheme by adopting the linear MAD prediction model that uses the actual MAD in the same position in the previous frame. M is used to allocate required bit budget to obtain the minimal quality for the non-target region. In our system, M is set to 3. After allocating the bit budget to the non-target region, the remaining bit budget for the current frame is allocated to the target region. The bit budget for the target region T_i^Ψ is obtained by the following:

$$T_i^\Psi = T_i - T_i^\Omega. \quad (10)$$

After obtaining the bit budget for the target region, the target SSIM for the current frame can be assigned according to the user-defined SSIM or according to the statistical results of the previous target region bitrate for a different SSIM. If the bit budget for the target region is sufficient to achieve the

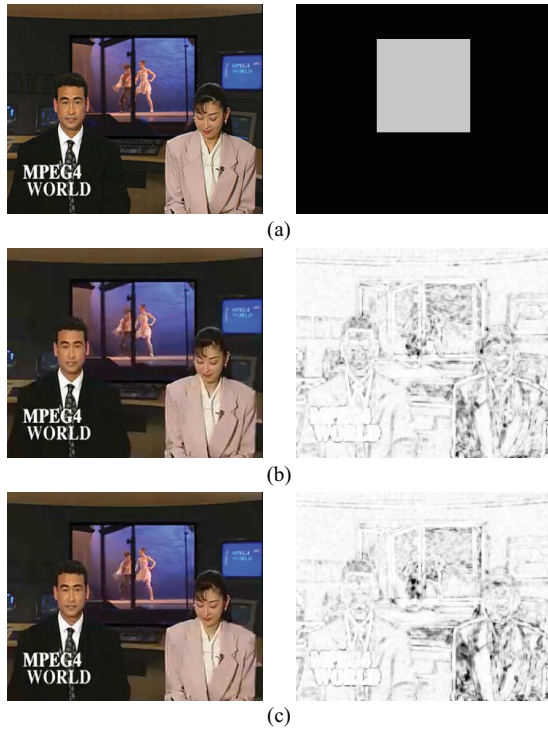


Fig. 14. Subjective demonstration of performance of the proposed system. The left side is the 88th frame of the reconstructed sequence “news.” The right side is the corresponding SSIM map. (a) Riginal frame and the target region map (white region). (b) JM RC Bitrate:2576 ROI SSIM:0.8965. (c) Proposed Bitrate:1880 ROI SSIM:0.9065.

user-defined SSIM, the user-defined SSIM becomes the target SSIM. Otherwise, if the user-defined SSIM is too high and requires a significant number of bits, an appropriate target SSIM must be estimated to avoid exceeding the bit budget. The estimated SSIM ($SSIM_i^\Psi$) for the target bits of the target region Ψ in the i th frame is estimated as follows:

$$SSIM_i^\Psi = \arg \min_{SSIM} |T_i^\Psi - T_{i-1}^{\Psi, SSIM}| \quad (11)$$

where $T_{i-1}^{\Psi, SSIM}$ is the total bits of the target region Ψ in the $(i-1)$ th frame to achieve a specific SSIM value. If the user-defined SSIM is smaller than $SSIM_i^\Psi$, that means T_i^Ψ is sufficient, and the user-defined SSIM becomes the target SSIM. Otherwise, $SSIM_i^\Psi$ becomes the target SSIM. For example, $T_{i-1}^{\Psi, SSIM=0.85}$ is the estimated bitrate for the target region of the $(i-1)$ th frame when the target SSIM is set to 0.85. If $T_{i-1}^{\Psi, SSIM=0.85}$ is closest to T_i^Ψ than any other $T_{i-1}^{\Psi, SSIM=0.80}$ to 0.95 , the estimated SSIM will be set to 0.85 for the current frame. In Sec. II-F, we will present the method to gather the statistical results of the previous target region bitrate T_{i-1}^Ψ .

D. Determine Target Bits and QP of the Current BU

At this stage, the target bit budget is assigned to the current BU using the similar scheme as in [6]. The different points are that the total bit budgets for target regions and non-target regions are determined in the previous stage. Moreover, only the BUs in the target region are used to determine the target bits for the current BU in the same target region. Let $T_{r,i}^\Psi$

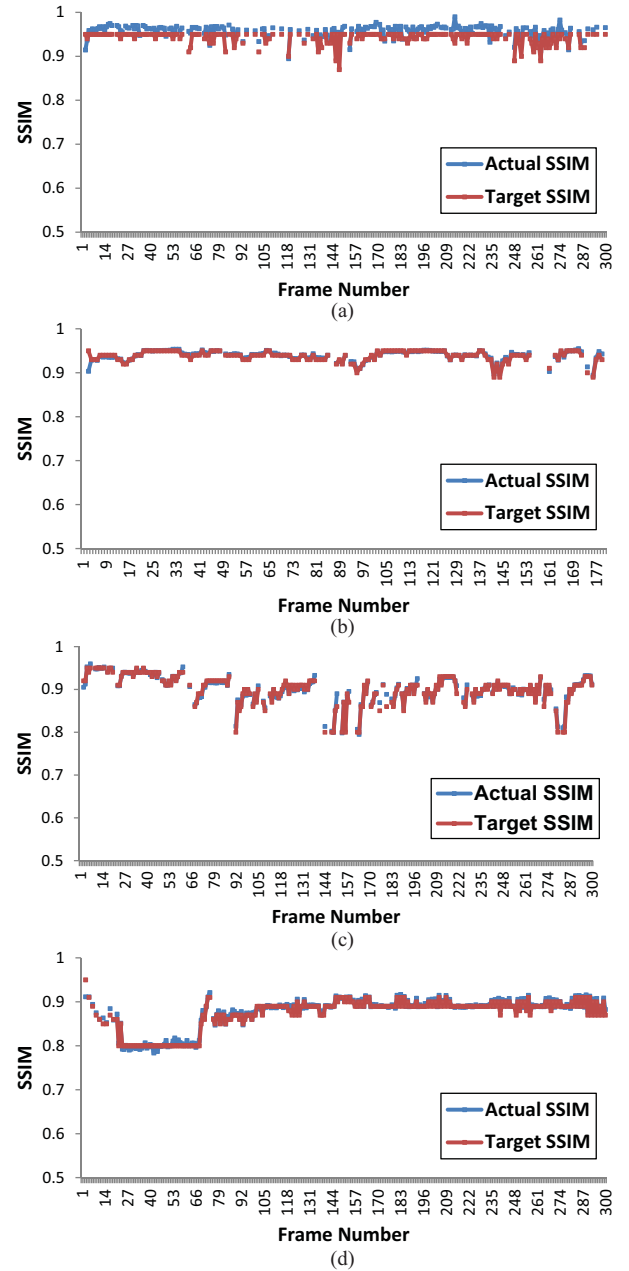


Fig. 15. Performance comparison of the target SSIMs and the actual SSIMs in the proposed system. (a) News. (b) Foreman. (c) Crew. (d) Coastguard.

denote the number of remaining bits for the remaining BUs of the target region in the i th picture. The target bits $t_{i,m}^\Psi$ for the m th BU of the target region in the i th picture are given as follows:

$$t_{i,m}^\Psi = T_{r,i}^\Psi \frac{MAD_{i,m}^\Psi}{\sum_{k=m}^{N_{unit}^\Psi} MAD_{i,k}^\Psi} \quad (12)$$

where N_{unit}^Ψ is the total number of BUs in the target region of the current frame and $MAD_{i,m}^\Psi$ is the MAD value of the m th BU in the target region in the i th frame. For the non-target regions, the target bit allocation method is similar to the method for target region BU, except the total bit budget is from T_i^Ω and only the BUs in the non-target regions are used to calculate the target bit budget for the current BU.

After the target bits for a BU is determined, we need to compute the corresponding QP. We adopted the quadratic R-Q model, which is also employed in the JVT-G012 due to its efficient performance. The QP is computed as follows:

$$t_{i,m}^{\Psi} = \left(b_1 \frac{MAD}{Q_{step}} + b_2 \frac{MAD}{Q_{step}^2} \right) \quad (13)$$

where b_1 and b_2 are model parameters and Q_{step} is the quantization step size, which has a one-to-one mapping relationship with QP. The details of above R-Q model can be found in [6].

E. Target Region BU QP Adjustment

The QP obtained before the mode decision stage is denoted as $QP_{initial}$. This stage is entered, only when processing the BUs in the target region. The BUs in the non-target region will use $QP_{initial}$ as the final QP to perform the remaining encoding. For the BUs in the target region, $SSIM_{pred}$, target SSIM (denoted as $SSIM_{rec}$ in SSIM-Q model), and var can be obtained after the mode decision. Those parameters can be used to evaluate the required QP to achieve $SSIM_{rec}$ based on the proposed SSIM-Q model. Note that the QP needs to be adjusted at this stage, only when the $SSIM_{pred}$ of the current BU is smaller than the predefined SSIM.

F. Model Update and Post-Processing

After the final QP has been decided, the normal encoding flows, such as transform, quantization and entropy, are performed to encode the current BU. The R-Q and MAD models need to be continually updated to adapt to the characteristics of the video content. In this work, we applied the same model update scheme as in [6] to update the model parameters.

In order to provide $T_i^{SSIM=0.80 \text{ to } 0.95}$ to estimate the target SSIM for the next frame, the bitrates for different target SSIMs have to be computed. The direct method is to encode the current BU using different target SSIMs, that is 0.80 to 0.95 in this work, and then collect the bitrate information after encoding the current BU. However, the computational complexity of the direct method is extremely high due to the multiple encoding procedures for different target SSIMs. To overcome this issue, we propose a prediction method of $T_i^{SSIM=0.80 \text{ to } 0.95}$ based on the R-Q model of (13) and the proposed SSIM-Q model of (8). Assuming we want to gather the statistics of a target SSIM 0.85, we use (8) to estimate the QP since we already know $\Delta SSIM_{rec}$, $SSIM_{pred}$, and var . Thereafter, we can apply (13) to estimate the bitrate of the current BU since we already have the QP and the MAD of the current BU. Applying the same method, we can predict the target bitrate for a specific target SSIM without encoding the current BU using some specific target SSIMs.

III. EXPERIMENTAL RESULTS

The schemes to be compared are implemented on the JM reference software version 14.0, for which all modules remain unchanged except the rate control. The experimental setting is as follows.

- 1) Baseline profile is used.
- 2) The first 300 frames of each test sequence are encoded.

TABLE II
PERFORMANCE OF THE PROPOSED SYSTEM WITH SUFFICIENT BITRATE BUDGET. THE SYSTEM ALLOCATES THE BITRATE TO THE ROI REGIONS TO ACHIEVE THE TARGET SSIM

Sequence	Target SSIM	Target Bitrate (kbits)	Bitrate (kbits)	SSIM	SSIM Error
Crew	0.85	1100	1100.3	0.8738	0.0238
	0.90	1100	1100.4	0.9069	0.0069
	0.95	1100	1100.4	0.9550	0.0050
	0.85	1500	1500.1	0.8823	0.0323
	0.90	1500	1500.1	0.9126	0.0126
	0.95	1500	1500.1	0.9556	0.0056
Coastguard	0.85	1100	1100.1	0.8525	0.0025
	0.90	1100	1100.1	0.9046	0.0046
	0.95	1100	1100.1	0.9112	-0.0388
	0.85	1500	1500.0	0.8520	0.0020
	0.90	1500	1500.0	0.9030	0.0030
	0.95	1500	1499.9	0.9621	0.0121
Paris	0.85	480	480.5	0.8856	0.0356
	0.90	480	480.2	0.9133	0.0133
	0.95	480	480.3	0.9537	0.0037
	0.85	600	600.3	0.8918	0.0418
	0.90	600	600.1	0.9173	0.0173
	0.95	600	600.3	0.9581	0.0081
News	0.85	480	480.7	0.8865	0.0365
	0.90	480	480.7	0.9094	0.0094
	0.95	480	480.6	0.9636	0.0136
	0.85	600	600.8	0.8770	0.0270
	0.90	600	600.8	0.9113	0.0113
	0.95	600	600.6	0.9686	0.0186
Foreman	0.85	480	481.2	0.8808	0.0308
	0.90	480	481.3	0.9103	0.0103
	0.95	480	481.1	0.9503	0.0003
	0.85	600	601.6	0.8871	0.0371
	0.90	600	601.4	0.9125	0.0125
	0.95	600	601.4	0.9514	0.0014
Avg	0.85	844	845	0.8769	0.0269
	0.90	844	845	0.9101	0.0101
	0.95	844	844	0.9530	0.0030

- 3) Rate distortion optimization is enabled.
- 4) Search range is 32.
- 5) Fast motion estimation EPZS is turned on.
- 6) GOP structure is IPPP.
- 7) One MB per BU is processed for CIF sequences.
- 8) QP 36 is taken as the initial QP for the first I frame and the first P frame.

To evaluate the proposed system, two cases are presented in this section. The first one has the sufficient bitrate budget to adjust quality to the target SSIM, while the other does not have sufficient bitrate budget. In the proposed system, the QP adjustment scheme adjusts QP to achieve the target quality according to the proposed SSIM-Q model under the bitrate constraint. When the target SSIM is too high, the QPs become extremely low to let the target regions to meet the target quality. In this situation, bitrate demand is increased hugely. So the first one case can evaluate the accuracy of the proposed SSIM-Q model and the proposed region-based rate control scheme. On the other hand, if the user-defined target SSIM can not be achieved under the bitrate constraint, the proposed scheme decreases the target SSIM adaptively to avoid the QP adjustment scheme consuming too many bitrate. the second case can evaluate the performance of the bit allocation of

TABLE III
PERFORMANCE OF THE DQW BIT ALLOCATION SCHEME WITH DIFFERENT WEIGHTING FACTORS FOR
QP CROPPING CASE AND NO QP CROPPING CASE

Weighting Factor	With QP Cropping						Without QP Cropping					
	Foreman		Paris		News		Foreman		Paris		News	
	bitrate (kbits)	SSIM	bitrate (kbits)	SSIM	bitrate (kbits)	SSIM	bitrate (kbits)	SSIM	bitrate (kbits)	SSIM	bitrate (kbits)	SSIM
0.1	478.4	0.9829	479.0	0.9576	477.0	0.9853	4695.5	0.9992	4902.3	0.9930	900.3	0.9991
0.2	478.4	0.9829	479.0	0.9576	478.4	0.9853	3174.1	0.9972	4659.1	0.9927	621.1	0.9975
0.3	478.4	0.9829	479.0	0.9576	478.8	0.9885	1839.1	0.9922	4761.8	0.9925	422.6	0.9945
0.4	478.4	0.9829	479.0	0.9576	479.2	0.9844	921.2	0.9806	5643.3	0.9936	395.5	0.9918
0.5	478.8	0.9829	479.0	0.9576	479.3	0.9843	480.3	0.9598	5417.2	0.9926	272.8	0.9767
0.6	478.1	0.9829	479.0	0.9576	480.2	0.9836	314.7	0.9439	8104.4	0.9978	480.0	0.9861
0.7	480.1	0.9599	479.0	0.9576	479.8	0.9835	192.7	0.9219	713.2	0.9689	480.4	0.9827
0.8	479.7	0.9594	479.0	0.9576	480.5	0.9826	475.3	0.8975	476.7	0.9678	480.4	0.9780
0.9	480.5	0.9580	479.8	0.9595	480.3	0.9824	479.6	0.9500	482.6	0.9564	480.6	0.9717
1.0	481.0	0.9530	480.9	0.9489	480.2	0.9800	480.3	0.9435	480.2	0.9493	480.3	0.9634
1.1	482.4	0.9417	479.5	0.9427	481.5	0.9749	481.5	0.9311	479.3	0.9386	481.4	0.9576
1.2	480.2	0.9381	479.9	0.9399	481.2	0.9727	660.0	0.9143	89.6	0.7584	481.6	0.9525
1.3	480.2	0.9381	479.9	0.9399	481.6	0.9710	2081.0	0.9452	137.3	0.7606	479.6	0.9081
1.4	480.2	0.9381	479.9	0.9399	481.6	0.9710	4027.9	0.9530	207.7	0.7695	477.4	0.8799
1.5	480.2	0.9381	479.9	0.9399	481.6	0.9710	5977.1	0.9596	309.3	0.7675	374.6	0.8806
1.6	480.2	0.9381	479.9	0.9399	481.6	0.9710	7529.9	0.9768	439.5	0.7660	451.7	0.8773
1.7	480.2	0.9381	479.9	0.9399	481.6	0.9710	8809.1	0.9828	661.4	0.7555	429.6	0.8800
1.8	480.2	0.9381	479.9	0.9399	481.6	0.9710	9258.7	0.9809	1026.0	0.7595	454.8	0.8712
1.9	480.2	0.9381	479.9	0.9399	481.6	0.9710	9814.0	0.9852	1438.5	0.7491	433.7	0.8729
2.0	480.2	0.9381	479.9	0.9399	481.6	0.9710	9778.5	0.9850	1814.1	0.7490	463.3	0.8692

the proposed region-based rate control scheme under bitrate constraint.

We compare our method with the traditional bit allocation methods that adjust QP for different video region according to some weighting factors, which is determined by some perceptual cue [14], [16], [17], [22]. Those methods try to allocate bit to different video regions to achieve the quality enhancement for some regions. To compare the performance of the proposed system, we take the method in [22] as a comparison case. In [22], they use visual attention cue as guidance maps to be the weighting factors to adjust the QPs of the encoding regions. The scheme to adjust QP for each location i is simplified as follows:

$$QP_i = W_i * Q_{baseline} \quad (14)$$

where $QP_{baseline}$ is the initial frame layer QP after the frame layer rate control scheme as described in [6], and w_i is the weight coefficient in area i . In this paper, we set each location i as one MB and denote the method as direct QP weighting method (DQW).

A. Quality-Regulable Video Coding System With Sufficient Bitrate Budget

To show the ability of quality scalability and rate control, we firstly show the drawback of the traditional DQW method. Table III shows the coding results of the actual bitrate and SSIM for the target regions with different weighting factors from 0.1 to 2.0. There is one case that adjusts QP with QP cropping scheme, that means the smallest QP is cropped to $QP_{baseline} - 2$, while the biggest QP is set to $QP_{baseline} + 3$. Another case is that the range of QP adjustment is not restricted unless it exceeds 51 or smaller than zero. From the

results, the QP cropping case shows that the ability of quality scalability is bad when the weighting factors become larger enough or smaller enough. In this situation, the coding results are the same because the QPs are cropped to a specified range. For example, the experimental results of “foreman” sequence show that when the weighting factor is smaller than 0.7 or larger than 1.3, the coding results are the same resulting in worse performance of the ability of quality scalability. For the other one case, we can see that the range of quality scalability is good because the QP is not restricted the range. As the DQW method adjusts QPs directly depending on the weighting factor and the rate-distortion is not considered in rate control scheme, they cause the QPs variation becomes large resulting in the poor performance of rate control if the weighting factor is too large or too small. For example, the experimental results of “foreman” sequence show that the coding bitrate becomes unconstrained and unpredictable when the weighting factor is smaller than 0.8 or larger than 1.1.

To evaluate the performance of the proposed system, we set the target SSIM for the ROI region to be constant for the entire sequence. Each BU in the ROI will adjust the QP to achieve the target SSIM according to the SSIM-Q model. Table II shows the accuracy of the proposed system for the test sequences with different target SSIMs and target bitrate. We set the target SSIM and compare it with the actual SSIM after coding. The average SSIM error for all sequences is 0.003, 0.0101, and 0.0269 for the target SSIMs of 0.95, 0.90, and 0.85 respectively. Moreover, the bitrate error is below 0.1%. Note that only when the $SSIM_{pred}$ of the MBs is smaller the target SSIM, the SSIM values of the MBs are considered as the comparison objects. Comparing the average SSIM of the whole target region is not necessary as the $SSIM_{pred}$ values of some MBs are larger than target SSIM. Our scheme tries to enhance the

TABLE IV
PERFORMANCE OF THE PROPOSED REGION-BASED PERCEPTUAL QUALITY-REGULABLE RATE CONTROL SCHEME

Sequence	JM [6]				DQW [22]			Proposed		
	Target Bit Rate (kbits)	Bit Rate (kbits)	ROI SSIM	Bitrate Error	Bit Rate (kbits)	ROI SSIM	Bitrate Error	Bit Rate (kbits)	ROI SSIM	Bitrate Error
Crew	192.0	192.2	0.7949	0.11%	191.5	0.8063	-0.28%	192.4	0.8107	0.21%
	288.0	288.3	0.8320	0.11%	287.1	0.8454	-0.30%	288.3	0.8491	0.11%
	384.0	384.3	0.8573	0.07%	382.9	0.8674	-0.29%	384.3	0.8739	0.09%
	480.0	480.3	0.8766	0.05%	479.0	0.8875	-0.20%	480.3	0.8926	0.07%
Coastguard	192.0	192.1	0.7855	0.03%	190.8	0.8093	-0.60%	192.2	0.8097	0.11%
	288.0	288.1	0.8207	0.03%	286.0	0.8558	-0.71%	288.2	0.8543	0.08%
	384.0	384.1	0.8484	0.03%	382.3	0.8845	-0.45%	384.4	0.8821	0.10%
	480.0	480.1	0.8706	0.03%	477.1	0.9040	-0.61%	480.2	0.9021	0.04%
Paris	96.0	96.2	0.8518	0.18%	96.2	0.8610	0.25%	101.8	0.8702	5.99%
	192.0	192.4	0.8943	0.22%	191.8	0.9036	-0.13%	193.6	0.9194	0.82%
	288.0	288.7	0.9161	0.23%	287.7	0.9302	-0.11%	288.9	0.9339	0.30%
	384.0	384.7	0.9338	0.18%	383.6	0.9409	-0.11%	384.7	0.9525	0.17%
News	96.0	96.3	0.8922	0.29%	96.1	0.9131	0.10%	96.5	0.9395	0.52%
	192.0	192.4	0.9427	0.21%	192.3	0.9514	0.13%	192.7	0.9623	0.34%
	288.0	288.6	0.9606	0.19%	288.4	0.9658	0.15%	288.4	0.9716	0.15%
	384.0	384.8	0.9703	0.20%	384.5	0.9738	0.13%	384.4	0.9775	0.10%
Foreman	96.0	96.3	0.8256	0.31%	96.1	0.8421	0.07%	97.7	0.8260	1.74%
	192.0	192.6	0.8880	0.31%	192.2	0.9041	0.11%	194.2	0.8989	1.14%
	288.0	289.0	0.9146	0.34%	287.5	0.9268	-0.18%	288.7	0.9251	0.25%
	384.0	385.2	0.9303	0.30%	384.1	0.9398	0.02%	384.8	0.9391	0.22%
Avg	134.4	134.6	0.8300	0.15%	134.1	0.8464	-0.19%	136.1	0.8512	1.27%
	230.4	230.8	0.8755	0.16%	229.9	0.8921	-0.23%	231.4	0.8968	0.43%
	326.4	326.9	0.8994	0.16%	325.7	0.9149	-0.20%	326.9	0.9173	0.17%
	422.4	423.0	0.9163	0.14%	421.7	0.9292	-0.18%	422.9	0.9328	0.11%

MBs that have smaller SSIMs than the target SSIM to achieve the values of the target SSIM. Fig. 9 shows the differences between the target SSIM value and the actual SSIM values for the adjusted region frame-by-frame. The SSIM values of the actual decoded video using the proposed system are close to the user-defined SSIM. However, the DQW method produces more fluctuant curves because of the poor ability of quality scalability. Note that the actual SSIM curve in Fig. 9(c) has more fluctuations than other sequences, and that the actual SSIM curve of the first few frames also has more fluctuation than that of the subsequent frames. The main reason is that we fixed the initial quantization parameter for all test sequences, which resulted in the initial frames having an actual SSIM value that is much higher than the target SSIM. Therefore, the inaccurate initial SSIM causes the subsequent frames to estimate the SSIM value poorly. Another reason is that the motion and scene of the test sequence “crew” in Fig. 9(c) is too high to estimate the quantization parameter well. For complicated sequences, a small quantization parameter change will cause the video quality to fluctuate wildly. These are the reasons why the SSIM curve in Fig. 9(c) has more fluctuations than other sequences, and why the SSIM curve of the first few frames is not as smooth as that of the subsequent frames. However, because the SSIM curve eventually becomes more stable, we know that the quality of the video encoded by the proposed system will converge towards the target SSIM quality.

Figs. 10 and 11 show the subjective video quality for different target SSIMs and the actual SSIMs after coding. The right side image is the SSIM map in which a whiter pixel

value indicates a better SSIM (closer to 1). We can see that the actual SSIM is close to the target SSIM. From the SSIM map, we also can see that if a target SSIM is larger, the target region achieves better SSIM quality. Moreover, the quality-regulability for different regions in the same video also can be seen in these two figures.

B. Quality-Regulable Video Coding System Without Sufficient Bitrate Budget

If the target SSIM is too high to achieve under the current bitrate budget constraint, it is important for a well-developed rate control algorithm to adjust the QP to avoid consuming too many bits. In this situation, the proposed system can adjust the target SSIM to adjust the QPs to meet the bitrate constraint. Table IV shows the comparisons of the performance of the proposed region-based rate control, DQW, and JM rate control schemes under the low bitrate constraint. To enhance the target regions, we set the “M” parameter to 1/3 to allocate more bits to those regions than the non-target regions. We can see the bitrate has been well-controlled under our scheme and the bitrate error is below 0.5% in average, which is close to the error rate of JM rate control. On the other hand, we can see that the target region achieves SSIM averages that are better than the JM rate control averages by a difference of up to 0.0133 to 0.0225. The bitrate error is below 0.29%. Fig. 15 shows the target SSIM and the actual reconstructed SSIM frame-by-frame. The figures show that the proposed region-based rate control scheme can adjust the target SSIM automatically to avoid exceeding the bit budget. Due to the proposed

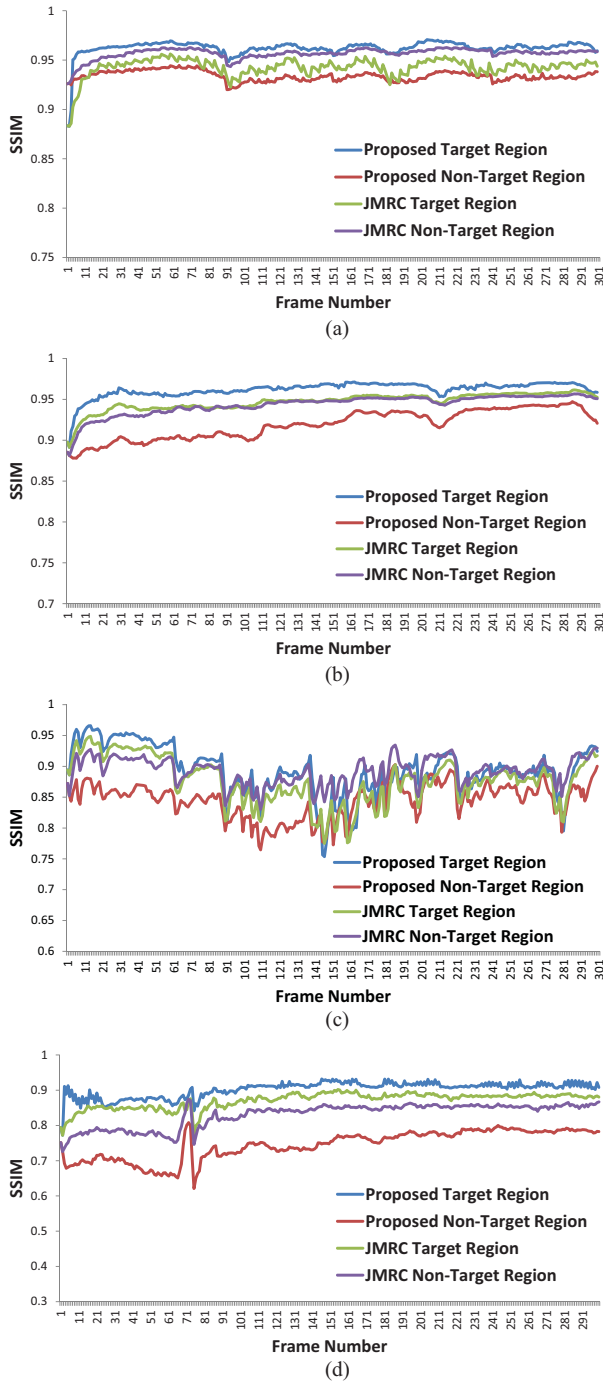


Fig. 16. Performance comparisons of the SSIM values of the target region and no-target region for the proposed region-based rate control system. (a) News. (b) Paris. (c) Crew. (d) Coastguard.

SSIM-Q model, the actual SSIM closely approximates the target SSIM. Fig. 16 shows the SSIM of target regions and non-target regions of each frame for various test sequences. It can be seen that, except for only a few frames, the proposed system can allocate a more accurate bit budget to the target region in order to achieve better SSIM quality.

To compare to the DQW rate control scheme, we set the weighting factor as 0.85 for all test sequences because the weighting factor is the value which can generate stable bitrate under the bitrate constraint. To encode video using this

weighting factor also can result in the quality enhancement of the target regions. From the experimental results, the proposed system can generate as stable bitrate as DQW method using the weighting factor. Moreover, the proposed system produces the better quality of SSIM. However, we should note that our method allocates bitrate to target region bounded by a parameter M as shown in (9). The DQW method produces the better quality of SSIM for the target regions by adjusting weighting factor. The results show that the proposed system can produce the better quality of SSIM in average while the DQW is under the situation of stable bitrate with a weighting factor 0.85. The comparison of subjective quality is shown in Fig. 12, Figs. 13 and 14. It can be seen that the proposed scheme generates better quality than the JM reference software with MB rate control. Figs. 13 and 14 show the performance of the proposed system when applied to different regions in the same video.

This evaluation is based on using the proposed quality-regulable coding system, instead of setting a constant target SSIM for each sequence. Fig. 9 shows the SSIM values for various test sequences. The target SSIM value for each frame is adjusted in order to meet the bit budget, and the difference between the actual SSIM after encoding and the target SSIM is very small.

IV. CONCLUSION

We propose a system design for a region-based perceptual quality-regulable H.264 video encoder. The proposed model analyzes the video signal and the information from the video coding loop to derive a more appropriate quantization parameter for the current coding MB. We adopted a structure similarity index as the quality metric for distortion-quantization modeling, and developed a bit allocation and rate control scheme for a perceptual quality-regulable video coding system. Compared to JM reference software with macroblock layer rate control, the proposed algorithm can effectively enhance perceptual quality for target video regions.

REFERENCES

- [1] Z. Chen and K. N. Ngan, "Recent advances in rate control for video coding," *Signal Process., Image Commun.*, vol. 22, no. 1, pp. 19–38, Jan. 2007.
- [2] T. Chiang and Y.-Q. Zhang, "A new rate control scheme using quadratic rate distortion model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 1, pp. 246–250, Feb. 1997.
- [3] H. J. Lee, T. Chiang, and Y.-Q. Zhang, "Scalable rate control for MPEG-4 video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 6, pp. 878–894, Sep. 2000.
- [4] W. G. S. Ma and Y. Lu, "Rate-distortion analysis for H.264/AVC video coding and its application to rate control," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 12, pp. 1533–1544, Dec. 2005.
- [5] Y. K. K. Z. He and S. K. Mitra, "Low-delay rate control for dct video coding via ρ -domain source modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 8, pp. 928–940, Aug. 2001.
- [6] Z. G. Li, F. Pan, K. P. Lim, G. Feng, X. Lin, and S. Rahardja, "Adaptive basic unit layer rate control for JVT," in *Proc. 7th Meeting JVT-G012-r1*, Pattaya II, Thailand, Mar. 2003.
- [7] M. Y. S. D. K. Kwon and C. C. J. Kuo, "Rate control for H.264 video with enhanced rate and distortion models," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 5, pp. 517–529, May 2007.
- [8] C. An and T. Q. Nguyen, "Iterative rate-distortion optimization of H.264 with constant bit rate constraint," *IEEE Trans. Image Process.*, vol. 17, no. 9, pp. 1605–1615, Sep. 2008.

- [9] J. Dong and N. Ling, "A context-adaptive prediction scheme for parameter estimation in H.264/AVC macroblock layer rate control," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 8, pp. 1108–1117, Aug. 2009.
- [10] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [11] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, Feb. 2006.
- [12] D. M. Chandler and S. S. Hemami, "VSNR: A wavelet-based visual signal-to-noise ratio for natural images," *IEEE Trans. Image Process.*, vol. 16, no. 9, pp. 2284–2298, Sep. 2007.
- [13] B. Wang, Z. Wang, Y. Liao, and X. Lin, "HVS-based structural similarity for image quality assessment," in *Proc. 9th Int. Conf. Signal Process.*, Oct. 2008, pp. 1194–1197.
- [14] X. Yang, W. Lin, Z. Lu, X. Lin, S. Rahardja, E. Ong, and S. Yao, "Rate control for videophone using local perceptual cues," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 4, pp. 496–507, Apr. 2005.
- [15] Y. Sun, I. Ahmad, D. Li, and Y. Q. Zhang, "Region-based rate control and bit allocation for wireless video transmission," *IEEE Trans. Multimedia*, vol. 8, no. 1, pp. 1–10, Feb. 2006.
- [16] C. W. Tang, "Spatiotemporal visual considerations for video coding," *IEEE Trans. Multimedia*, vol. 9, no. 2, pp. 231–238, Feb. 2007.
- [17] Y. Liu, Z. G. Li, and Y. C. Soh, "Region-of-interest based resource allocation for conversational video communication of H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 1, pp. 134–139, Jan. 2008.
- [18] Z. Chen and C. Guillemot, "Perceptually-friendly H.264/AVC video coding based on foveated just-noticeable-distortion model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 6, pp. 806–819, Jun. 2010.
- [19] G. L. Wu, T. H. Wu, Y. J. Fu, and S. Y. Chien, "Perception-aware H.264/AVC encoder with hardware perception analysis engine," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2010, pp. 790–795.
- [20] G. L. Wu, T. H. Wu, and S. Y. Chien, "Algorithm and architecture design of perception engine for video coding applications," *IEEE Trans. Multimedia*, vol. 13, no. 6, pp. 1181–1194, Dec. 2011.
- [21] T. S. Ou, Y. H. Huang, and H. H. Chen, "SSIM-based perceptual rate control for video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 5, pp. 682–691, May 2011.
- [22] Z. Li, S. Qin, and L. Itti, "Visual attention guided bit allocation in video compression," *Image Vis. Comput.*, vol. 29, no. 1, pp. 1–14, Jan. 2011.
- [23] G. L. Wu, Y. J. Fu, and S. Y. Chien, "System design of perceptual quality-regulable H.264 video encoder," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2012, pp. 509–514.
- [24] G. L. Wu, Y. J. Fu, and S. Y. Chien, "Region-based perceptual quality regulable bit allocation and rate control for video coding applications," in *Proc. Visual Commun. Image Process.*, Nov. 2012, pp. 1–6.
- [25] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3440–3451, Nov. 2006.
- [26] Z. Wang, Q. Li, and X. Shang, "Perceptual image coding based on a maximum of minimal structural similarity criterion," in *Proc. IEEE Int. Conf. Image Process.*, Sep.–Oct. 2007, pp. 121–124.
- [27] T. Richter and K. J. Kim, "A MS-SSIM optimal JPEG 2000 encoder," in *Proc. Data Compress. Conf.*, 2009, pp. 401–410.
- [28] Y. H. Huang, T. S. Ou, P. Y. Su, and H. H. Chen, "Perceptual rate-distortion optimization using structural similarity index as quality metric," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 11, pp. 1614–1624, Nov. 2010.
- [29] S. S. Channappayya, A. C. Bovik, C. Caramanis, and R. W. Heath, "Design of linear equalizers optimized for the structural similarity index," *IEEE Trans. Image Process.*, vol. 17, no. 6, pp. 857–872, Jun. 2008.
- [30] J. L. Devore and N. R. Farnum, *Applied Statistics for Engineers and Scientists*. Pacific Grove, CA, USA: Duxbury Press, 1999.



Guan-Lin Wu received the B.S. degree from the Department of Electrical Engineering, National Cheng-Kung University, Tainan, Taiwan, in 2003, and the M.S. and Ph.D. degrees from the Graduate Institute of Electronics Engineering, National Taiwan University (NTU), Taipei, Taiwan, in 2005 and 2010, respectively.

He is currently a Post-Doctoral Researcher with the Media IC and System Laboratory, Graduate Institute of Electronics Engineering, NTU. His current research interests include algorithms and very large scale integration architectures of video signal processing, reconfigurable computing, and system-on-a-chip architecture design.

Yu-Jie Fu received the B.S. and M.S. degrees from the Department of Electrical Engineering, National Taiwan University (NTU), Taipei, Taiwan, in 2009 and 2011, respectively.

He was with the Media IC and System Laboratory, Graduate Institute of Electronics Engineering, NTU, from 2009 to 2011. His current research interests include video coding technology, image processing, and very large scale integration implementation.



Sheng-Chieh Huang was born in Chunghua, Taiwan, in 1967. He received the B.S. degree in hydraulic ocean engineering from Nation Cheng Kung University, Tainan, Taiwan, in 1991, and the M.S. and Ph.D. degrees in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1993 and 1999, respectively.

He is currently an Assistant Professor with the Department of Electrical Engineering, National Chiao-Tung University, Hsinchu, Taiwan. His current research interests include very large scale integration design in DSP/DIP architecture design, video coding system, and Traditional Chinese Medicine system-on-a-chip design.



Shao-Yi Chien (S'99–M'04) received the B.S. and Ph.D. degrees from the Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan, in 1999 and 2003, respectively.

He is currently a Professor with the Graduate Institute of Electronics Engineering and Department of Electrical Engineering, National Taiwan University, where he joined as an Assistant Professor in 2004. From 2003 to 2004, he was a Research Staff with Quanta Research Institute, Tao Yuan County, Taiwan. He has authored or co-authored more than 190 papers in journals and conferences. His current research interests include video segmentation algorithm, intelligent video coding technology, perceptual coding technology, image processing for digital still cameras and display devices, computer graphics, and the associated very large scale integration and processor architectures.

Dr. Chien is an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—I: REGULAR PAPERS, and *Circuits, Systems and Signal Processing* (Springer). He was a Guest Editor for the *Journal of Signal Processing Systems* (Springer) in 2008. He is on the technical program committees of several conferences, such as ISCAS, ICME, SiPS, A-SSCC, and VLSI-DAT.